Thesis for the Degree of Doctor of Philosophy

Energy-Efficient Digital Signal Processing for Fiber-Optic Communication Systems

Christoffer Fougstedt

CHALMERS

Department of Computer Science and Engineering
Chalmers University of Technology
Göteborg, Sweden, 2019

Energy-Efficient Digital Signal Processing for Fiber-Optic Communication Systems
Christoffer Fougstedt

Göteborg, Sweden, 2019

ISBN: 978-91-7905-165-5

Doktorsavhandlingar vid Chalmers Tekniska Högskola
Ny serie Nr 4632
ISSN 0346-718X

Technical report 175D
Department of Computer Science and Engineering
Chalmers University of Technology
SE–412 96 Göteborg
Sweden
Telephone: +46–(0)31–772 10 00

Printed by Reproservice
Chalmers Tekniska Högskola
Göteborg, Sweden, 2019

High-throughput power-efficient DSP for fiber-optic communication Systems
Christoffer Fougstedt
Department of Computer Science and Engineering
Chalmers University of Technology

Abstract

Modern fiber-optic communication systems rely on complex digital signal processing (DSP) and forward error correction (FEC), which contribute to a significant amount of the over-all link power dissipation. Bandwidth demands are evergrowing and circuit technology scaling will due to fundamental reasons come to an end; energy-efficient design of DSP is thus necessary both from a sustainability perspective and a technical perspective. This thesis explores energy-efficient design of the sub-systems that are estimated to contribute to the majority of the receiver application-specific integrated-circuit power dissipation: chromatic-dispersion compensation, dynamic equalization, nonlinearity mitigation, and forward error correction. With a focus on real-time-processing circuit implementation of the considered algorithms, aspects such as finite-precision effects, pipelining, and parallel processing are explored, the impact on compensation and correction performance is investigated, and energy-efficient circuit implementations are developed. The sub-systems are investigated both individually, and in a system context. DSP designs showing significant energy-efficiency improvements are presented, as well as very high-throughput, energy-efficient, FEC designs. The subsystems are also considered in the context of datacenter interconnect links, and it is shown that DSP-based coherent systems are feasible even in power constrained settings.

Keywords: Application Specific Integrated Circuits, Communication Systems, Digital Signal Processing, Fiber Optic Communication, Non-linear Impairment Mitigation, Forward Error Correction

Publications

This thesis is based on the work contained in the following papers:

[A] **Christoffer Fougstedt**, Alireza Sheikh, Pontus Johannisson, and Per Larsson-Edefors. "Filter Implementation for Power-Efficient Chromatic Dispersion Compensation", *IEEE Photonics Journal*, **7202919**, Jun 2018.

[B] **Christoffer Fougstedt**, Pontus Johannisson, Lars Svensson, and Per Larsson-Edefors. "Dynamic Equalizer Power Dissipation Optimization", *Optical Fiber Communications Conference, OFC 2016*,

[C] **Christoffer Fougstedt**, Mikael Mazur, Lars Svensson, Henrik Eliasson, Magnus Karlsson, and Per Larsson-Edefors. "Time-Domain Digital Back Propagation: Algorithm and Finite-Precision Implementation Aspects", *Optical Fiber Communications Conference, OFC 2017*,

[D] **Christoffer Fougstedt**, Lars Svensson, Mikael Mazur, Magnus Karlsson, and Per Larsson-Edefors. "Finite-Precision Optimization of Time-Domain Digital Back Propagation by Inter-Symbol Interference Minimization", *Proceedings of 43rd European Conference and Exhibition on Optical Communications, ECOC 2017*,

[E] **Christoffer Fougstedt**, Lars Svensson, Mikael Mazur, Magnus Karlsson, and Per Larsson-Edefors. "ASIC Implementation of Time-Domain Digital Back Propagation for Coherent Receivers", *IEEE Photonics Technology Letters*, **30**, 13, 1179–1182, Jul 2018.

[F] **Christoffer Fougstedt**, Christian Häger, Lars Svensson, Henry D. Pfister, and Per Larsson-Edefors. "ASIC Implementation of Time-Domain Digital Backpropagation with Deep-Learned Chromatic Dispersion Filters", *Proceedings of 44st European Conference and Exhibition on Optical Communications, ECOC 2018.*

[G] **Christoffer Fougstedt**, Krzysztof Szczerba and Per Larsson-Edefors. "Low-Power Low-Latency BCH Decoders for Energy-Efficient Optical Interconnects", *Journal of Lightwave Technology*, **35**, 23, 5210–5207, Dec 2017.

[H] **Christoffer Fougstedt** and Per Larsson-Edefors. "Energy-Efficient High-Throughput VLSI Architectures for Product-Like Codes", *Journal of Lightwave Technology (top-scored)*, **37**, 2, 477–485, Jan 2019.

[I] **Christoffer Fougstedt**, Alireza Sheikh, Alexandre Graell i Amat, Gianluigi Liva, and Per Larsson-Edefors. "Energy-Efficient Soft-Assisted Product Decoders", *Optical Fiber Communications Conference, OFC 2019.*

[J] **Christoffer Fougstedt**, Oscar Gustafsson, Cheolyong Bae, Erik Börjeson, and Per Larsson-Edefors. "DSP and FEC Power Dissipation in 400G Coherent Data Center Interconnects",
Manuscript.

Related work by the author (not included in this thesis):

[K] **Christoffer Fougstedt**, Alireza Sheikh, Pontus Johannisson, Alexandre Graell i Amat, and Per Larsson-Edefors. "Power-Efficient Time-Domain Dispersion Compensation Using Optimized FIR Filter Implementation", *Signal Processing in Photonics Communications, SPPCom 2015.*

[L] Alireza Sheikh, **Christoffer Fougstedt**, Alexandre Graell i Amat, Pontus Johannisson, Per Larsson-Edefors, and Magnus Karlsson. "Dispersion Compensation Filter Design Optimized for Robustness and Power Efficiency", *Signal Processing in Photonics Communications, SPPCom 2015.*

[M] Krzysztof Szczerba, **Christoffer Fougstedt**, Per Larsson-Edefors, Petter Westbergh, Alexandre Graell i Amat, Lars Svensson, Magnus Karlsson, Anders Larsson, and Peter Andrekson. "Impact of Forward Error Correction on Energy Consumption of VCSEL-based Transmitters", *41st European Conference on Optical Communication, ECOC 2015.*

[N] Alireza Sheikh, **Christoffer Fougstedt**, Alexandre Graell i Amat, Pontus Johannisson, Per Larsson-Edefors, and Magnus Karlsson. "Dispersion Compensation FIR Filter with Improved Robustness to Coefficient Quantization Error", *Journal of Lightwave Technology*, **34**, 22, 5110–5117, Aug 2016.

[O] Lars Lundberg, **Christoffer Fougstedt**, Per Larsson-Edefors, Peter Andrekson, and Magnus Karlsson. "Power Consumption of a Minimal-DSP Coherent Link with a Polarization Multiplexed Pilot-Tone", *42nd European Conference on Optical Communication, ECOC 2016.*

[P] **Christoffer Fougstedt** and Per Larsson-Edefors. "Energy-Efficient High-Throughput Staircase Decoders", *Optical Fiber Communications Conference, OFC 2018.*

[Q] Per Larsson-Edefors, **Christoffer Fougstedt** and Kevin Cushon. "Implementation Challenges for Energy-Efficient Error Correction in Optical Communication Systems", *Signal Processing in Photonics Communications, SPPCom 2018*

[R] Lars Lundberg, Erik Börjeson, **Christoffer Fougstedt**, Mikael Mazur, Magnus Karlsson, Peter Andrekson, and Per Larsson-Edefors. "Power Consumption Savings Through Joint Carrier Recovery for Spectral and Spatial Superchannels", *44th European Conference and Exhibition on Optical Communications, ECOC 2018*

[S] Erik J Ryman, **Christoffer Fougstedt**, Lars Svensson, and Per Larsson-Edefors. "Custom versus Cell-Based ASIC Design for Many-Channel Correlators", *IEEE Workshop on Signal Processing Systems, IEEE SiPS 2018*

[T] Erik Börjeson, **Christoffer Fougstedt**, and Per Larsson-Edefors. "ASIC Design Exploration of Phase Recovery Algorithms for M-QAM Fiber-Optic Systems", *Optical Fiber Communications Conference, OFC 2019.*

[U] Erik Börjeson, **Christoffer Fougstedt**, and Per Larsson-Edefors. "Towards FPGA Emulation of Fiber-Optic Channels for Deep-BER Evaluation of DSP Implementations", *Signal Processing in Photonics Communications, SPPCom 2019*

Contents

Acknowledgement

First and foremost, I want to express my deepest gratitude to my main supervisor, Prof. Per Larsson-Edefors. I have not only had the privilege to enjoy his great support, generous sharing of knowledge, and his guiding throughout this endeavor, but I have also had the pleasure to enjoy interesting discussions, good music, literature, and film tips, and in general fun times both in office and abroad during travel. I am forever grateful.

I want to thank my co-supervisor Dr. Lars Svensson for his support, inspiring discussions and ideas, and in-general fun and interesting discussions on a plethora of work- and non-work-related subjects.

I want to thank my former co-supervisor Dr. Pontus Johannison for his support and guidance, my co-supervisor Prof. Magnus Karlsson for all good discussions (especially during conferences, which have truly helped in broadening my knowledge) and his support, and Prof. Jan Jonsson for his support.

During my time as a Ph.D. student, I have had the privilege to enjoy several wonderful collaborations, both in project, cross projects, and also cross universities. I have worked in projects which have been great team efforts, and I am very grateful for all individual contributions. I want to thank Mikael Mazur and Dr. Lars Lundberg, for inspiring collaborations, discussions, contributions, and good times in general. I want to thank Dr. Oscar Gustafsson and Cheolyong Bae for inspiring discussions and their contribution to our joint work. I want to thank Dr. Christian Häger for interesting and inspiring discussions, and his contributions on nonlinearity mitigation and machine learning. I want to thank Dr. Krzysztof Szczerba for sharing his knowledge, interesting discussion, contributions, and in-general inspiring positivity. I want to thank Dr. Alireza Sheikh and Dr. Henrik Eliasson for good discussions and contributions. I want to express my gratitude to my office mates, Dr. Erik Ryman, Dr. Kevin Cushon, Victor Åberg, Erik Börjeson, for all good discussions, fun times, research ideas, collaborations, and coffee. Their contributions have been invaluable, especially during tape-out crunches.

I also want to thank Prof. Alexandre Graell i Amat, Prof. Peter Andrekson, Prof. Erik Agrell, Dr. Jochen Schröder, the Computer Engineering Division, and the FORCE center at Chalmers. I want to thank the Knut and Alice Wallenberg foundation for financial support.

Finally, I want to express my deepest gratitude to my wonderful family, Dan, Carina, Andreas, and my love Klara, for their support throughout this journey. I love you.

Göteborg, 2019

Acronyms

ADC	analog-to-digital converter
ASIC	application-specific integrated circuit
AWGN	additive white Gaussian noise
BCH	Bose-Chaudhuri-Hocquenghem
BD	bounded-distance
BER	bit-error rate
BI-AWGN	binary-input additive white Gaussian noise
BPSK	binary phase-shift keying
CD	chromatic dispersion
CDC	chromatic-dispersion compensation
CMA	constant modulus algorithm
CMOS	complimentary metal oxide semiconductor
CPE	carrier-phase estimation
DBP	digital back propagation
DCF	dispersion-compensating fiber
DCI	data-center interconnect
DSP	digital signal processing
EDFA	erbium-doped fiber amplifier
FD-SOI	fully-depleted silicon-on-insulator
FEC	forward error correction
FFT	fast Fourier transform
FIR	finite impulse response
HD	hard decision
HDL	hardware description language
HPC	high-performance computing
iBDD	iterative bounded-distance decoding
IFFT	inverse fast Fourier transform
IIR	infinite impulse response
IM/DD	intensity-modulation direct detect
IS	impulse response
ISI	inter-symbol interference
KES	key-equation solver
LDPC	low-density parity check

LFSR	linear-feedback shift register
LLR	log-likelihood ratio
LO	local oscillator
LS-CO	least-squares constrained-optimization
LS-FB	least-squares full-band
LUT	lookup table
MD	minimum-distance
MIMO	multiple-input multiple-output
NCG	net coding gain
OH	overhead
OMA	optical modulation amplitude
OOK	on-off keying
OS	overlap-save
PAM	pulse-amplitude modulation
PM	polarization multiplexed
PMD	polarization-mode dispersion
QAM	quadrature-amplitude modulation
QPSK	quadrature phase-shift keying
RRC	root-raised cosine
RS	Reed-Solomon
SD	soft decision
SIR	signal-to-interference ratio
SMF	single-mode fiber
SPS, SaPS	samples per symbol
SQNR	signal-to-quantization-noise ratio
SR	scaled reliability
SSFM	split-step Fourier method
StPS	steps per span
TD	time-domain
VCSEL	vertical-cavity surface-emitting laser
VHDL	very high speed integrated circuit hardware description language
VLSI	very-large-scale integration
WDM	wavelength-division multiplexing

Chapter 1

Introduction

While digital signal processing (DSP) and forward error correction (FEC) are corner stones in enabling modern high-throughput fiber-optic communication, DSP and FEC are estimated to contribute a significant amount of overall link energy dissipation [1]. This thesis shows that implementation-focused design of algorithms and architectures can significantly reduce DSP and FEC energy dissipation.

Fiber-optic communication has come a long way since the invention of the low-loss fiber in 1970 [2], and nowadays long-haul networks span the globe. Transmitting data using light is conceptually simple: in its bare essence, we encode ones as high amplitude and zeros as low, and detect the received light. However, modern systems employ complex processing at high speed and put stringent requirements on system design. Coherent intradyne systems [3] revolutionized fiber-optic communication, and are in many ways more similar to typical wireless transmission systems rather than traditional on-off-keying-based optical communication. Coherent systems allow for capture of the full electromagnetical field envelope, and has thus enabled transmission of data on both phase and amplitude, as well as effective compensation of impairments. While traditionally mainly employed in long-haul systems, coherent technologies are expected to find its way into shorter-reach links as well [4, 5].

The fiber-optic communication channel can provide an immense bandwidth, which needs to be used as efficiently as possible in order to meet the ever-increasing demands of the modern digital society. While there is a significant scientific strive for pushing the limits in both reach and system throughput, metrics which are relatively straight-forward to quantify, increasing the throughput is only one side of the coin. Energy, or power efficiency aspects are equally or perhaps, depending on types of links, even more important as traditional performance metrics, but also typically more difficult to quantify, as we are now exploring scenarios where trade-offs come into play. Additionally, feature-size scaling in integrated circuit is—due to both fundamental limits and economical aspects—coming to an end [6], and future circuit technology may

no longer hold a promise for enabling implementation of increasingly advanced algorithms.

Although energy and power efficiency are related, energy and power efficiency requirements has a rather varied impact and needs to be differentiated. For example, on a micro scale, power efficiency is a key aspects due to packaging requirements, and thus of utmost concern in short range links such as data center interconnects, which are typically rather densely packed. On the other hand, energy efficiency is important on a macro scale; bandwidth demands of communication systems are rapidly rising, and energy efficiency considerations are thus of utmost importance in order to allow for sustainable development of the modern digital society.

In this thesis, we explore DSP and FEC algorithms and implementations which rather than pushing a single boundary, such as reach or throughput, strives towards achieving a trade-off between reach, throughput, and energy efficiency. Whereas both DSP and FEC algorithms are typically integrated in a single chip using digital logic, their behavior and thus implementation considerations may be vastly different. DSP algorithms typically operate in a stream-processing fashion, while commonly considered FEC schemes consists of both error detection and correction, where the correction logic is only used if errors are found. Essentially, DSP algorithms can be considered a part of the communication channel, which attempts to counteract systematic impairments, whereas FEC acts on message decisions, and attempts to counteract random errors in the transmitted information stream. Finally, we also explore system-level power dissipation aspects with a focus on coherent datacenter interconnect (DCI) links.

1.1 Thesis Outline

This thesis is structured as follows. Ch. 2 provides an introduction to channel models in general, and fiber-optic channel models in particular. The models and the approaches used in this work are discussed. Ch. 3 explores design and implementation of digital signal processing and forward error correction for fiber-optic communication, with a focus on modern semi-custom application-specific integrated-circuit design. Ch. 4 discusses fiber-optic communication systems with a focus on energy and power dissipation, focusing on the sub-systems relevant to this work. Ch. 5 summarizes the contributions of the included papers, and provides a future outlook. Finally, the included papers (paper A–J) are presented.

Chapter 2

Fiber-Optic Communication

In its bare essence, fiber optic communication encodes data on light, which is propagated through an optical fiber and detected at the receiving end. While the principle may sound simple, modern fiber optic communication systems rely on several complex subsystems in order to achieve reliable, high-throughput communication. Thus, we need to approach system design using a divide-and-conquer approach, and system models are therefore necessary. This chapter provides an introduction to communication channel models in general, and fiber-optic communication in particular, focusing on aspects relevant to this thesis. There are a wide variety of effects that impact system performance, including, but not limited to, optical impairments such as chromatic dispersion (CD), electrical impairments such as receiver bandwidth limitations, and non-idealities in signal processing. In addition, fiber optic communication systems often operate using wavelength-division multiplexing (WDM), and there are several cross-channel impairments that affect the system performance. However, the focus of this thesis is implementation aspects of DSP and FEC; we thus mainly concern ourselves with single-channel and DSP impairments.

2.1 Communication Channels

In the 1940's, Claude Shannon showed, using a general communication channel model, that an arbitrarily low error probability is achievable—even though the transmitted data is corrupted by noise—as long as the rate of transmission does not exceed the channel capacity [7]. However, he did not show how to practically approach this capacity bound. While the communication systems of today are vastly more advanced than those of concern in the early works, the models and ideas remain relevant to this day.

We want to transmit information from point A (the source) to point B (the destination), we assume that the information bitstream is uniformly distributed (and thus need not to concern ourselves with source coding). In order to do this, we need a transmitter (mapping data to a signaling scheme), a medium to communicate over, and a receiver (demapping signals to data). At the days of the foundation of communication theory, the transistor was in its infancy, and tubes were mainstream technology; DSP implementation aspects were thus for obvious reasons not taken into account. While we typically consider DSP a part of the receiver (and transmitter) system, effects such as for example rounding errors and non-ideal algorithm implementation adds noise and impairments, and can thus, from a demodulation and error correction viewpoint, be seen as a part of the communication channel. In this thesis, we focus on the receiver DSP and FEC.

Since the systems are complex, communication system models are required in order to allow for effective design and analysis of systems and algorithms. Since the focus here are implementation of DSP and FEC, we are interested in simple models that allows us to single out effects that appear in our implementations. Focusing on FEC, two useful, yet very simple, channel models are the binary-symmetric channel (BSC) and the binary-input additive white Gaussian noise channel (BI-AWGN), which are useful in designing and evaluating FEC schemes for hard-decision decoding and soft-decision decoding, respectively, assuming that long interleavers are used to decorrelate remaining impairments with temporal memory.

In the BSC, binary data is transmitted with a certain random, memoryless crossover probability, p. The probability of a transmitted bit being corrects is thus $1 - p$. The channel capacity is the maximum of the mutual information (the amount of information of one random variable that can be obtained by the observation of another) of the channel input and output, over all possible input distributions; in the case of BSC, the capacity is [8, Ch. 1]

$$C_{\mathrm{BSC}} = 1 - \mathrm{H}(p) \tag{2.1}$$

bits per channel use, where $\mathrm{H}(p)$ is the binary entropy function, and p is the crossover probability. Capacity is zero if $p = 0.5$. At other bit-error probabilities, error-free communication is theoretically possible, albeit at low rates. In practice, with realistic block lengths, we need to operate at a margin from the capacity limit (often referred to as the hard-decision Shannon limit); however, modern codes such as staircase codes [9] can approach the limit rather closely. For example, a R=0.94 staircase code can provide essentially error-free operation at a bit-error rate of $4.7 \cdot 10^{-3}$ [10]; the corresponding channel capacity at this BER is 0.96.

The BI-AWGN channel accepts binary inputs coded as +1 or -1, (commonly referred to as binary phase-shift keying (BPSK)), and outputs the transmitted value plus the added noise. Here, we can obtain both the estimated bits and per-bit reliabilities; since we have more information, intuitively we should be able to improve performance and obtain a higher capacity, which is indeed

the case [8, Ch. 1] . In the case of the BI-AWGN channel, we cannot obtain a closed-form expression, and calculating the capacity requires numerical integration [8, Ch. 1].

Although simplistic, the mentioned channel models are nevertheless useful in designing FEC for fiber-optic systems, as the BSC and the BI-AWGN capacities give the upper bound for decoders operating under the memoryless channel assumption using BPSK (or quarternary phase-shift keying (QPSK), as I and Q are orthogonal), and hard-decision decoding or soft-decision decoding, respectively. In the case of coherent dispersion-unmanaged links, the Gaussian noise model [11] has been shown to be rather accurate [12, 13]; the output from a link operating in the pseudo-linear regime, with properly designed DSP chain, and long pseudo-random interleaving can thus be assumed to behave similar to the simple models. It should be noted that this approach is suboptimal, since the capacity of a channel with memory is higher than the interleaved channel [14]. However, such code design is not considered in this thesis.

In the code rates of interest for current fiber-optic communication systems (R=0.7–0.95), soft-decision decoding can theoretically achieve approximately 1.1–1.5 dB of additional coding gain in comparison to hard-decision decoding [8, Ch. 1]. So why consider hard-decision decoding? In this thesis, for two main reasons: some simple fiber-optic systems, such as direct detection optical interconnects, do not allow for capture of soft information, and hard-decision decoding is less complex and lends itself well to high-throughput, low-power implementations.

2.1.1 The Fiber-Optic Channel

While simple channel models are useful in designing FEC schemes, such approaches rely on proper compensation of impairments in the fiber. Thus, more elaborate channel models are required in the design and evaluation of DSP algorithms. In general, experimental approaches are desirable to evaluate real-world performance. However, an experimental setup is essentially an all or nothing approach: it is in many cases not possible to single out the impact of a certain design parameter or choice. Here, we focus on single-channel impairments as we want to isolate DSP effects, and current (and, regarding circuits, for the foreseeable future) technology does not allow for integration of full WDM processing in real time. The main focus of this thesis is coherent systems, which allows for capture of the full electromagnetic field envelope and thus the use of advanced phase-amplitude modulation formats and digital signal processing. In addition, coherent receivers widely employ polarization-diverse transmission and thus 4-dimensional modulation (although often treated as 2×2-dimensional modulation). Fig. 2.1 shows an example of a DSP-based coherent receiver front-end. The input is split into two polarizations using a polarization beam splitter, and is mixed with a local oscillator laser (LO) and separated into in-phase and quadrature components using a 90° hybrid. The light is then detected and digi-

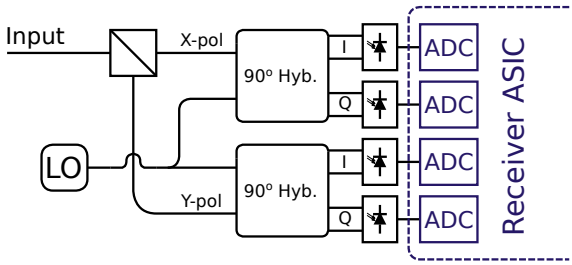


Figure 2.1: Block diagram of a coherent DSP-based receiver.

tized using photodiodes and analog-to-digital converters (ADC). Fig. 2.2 shows constellation diagrams of commonly employed modulation formats, where on-off keying is typically employed in traditional one-polarization intensity modulation direct-detect (IM/DD) systems while QPSK or 16-quadrature amplitude modulation (QAM) is often employed in both x- and y-polarization in coherent systems. The considered systems commonly operate in the C-band region (centered approximately around a wavelength of 1550 nm); at this region, fibers with a loss as low as 0.2 dB/km has been available since the end of the 1970's [15].

In the single-channel, single-polarization case, the fiber-optic channel can be modeled using the nonlinear Schrödinger equation [16, Ch. 2]:

$$\frac{\partial A}{\partial z} = (\hat{D} + \hat{N})A = \left(-\frac{j\beta_2}{2}\frac{\partial^2}{\partial t^2} - \frac{\alpha}{2}\right)A + j\gamma\left|A\right|^2 A, \tag{2.2}$$

where β_2 is the group-velocity dispersion parameter, α is the attenuation of the fiber, and γ is the nonlinear coefficient. If a polarization-multiplexed system is considered, the equation can be modified into the Manakov equation [17] which takes into account the power in both polarizations; nevertheless, the underlying behavior is similar and, for the sake of clarity, we will focus on the single-channel case. The equations consists of two major parts, the linear ($\hat{D}$) and the nonlinear ($\hat{N}$) parts, where the linear parts include the effect of chromatic dispersion (first term) and fiber attenuation (second term), while the nonlinear part models the nonlinear phase shift. We cannot obtain a closed-form solution of the equations, and we need to resort to numerical methods for solving. A common approach is to assume that, given short enough propagation, the linear and nonlinear parts can be assumed to act independently. We here simulate propagation by slicing the fiber into very small linear steps, with nonlinear operations intertwined. The linear steps are solved in the frequency domain, while the nonlinear steps are solved in the time domain; this is referred to as the split-step Fourier method [16, Ch. 2].

Chromatic dispersion, caused by the wavelength dependency of the fiber refractive index and thus different propagation speeds for the spectral compo-

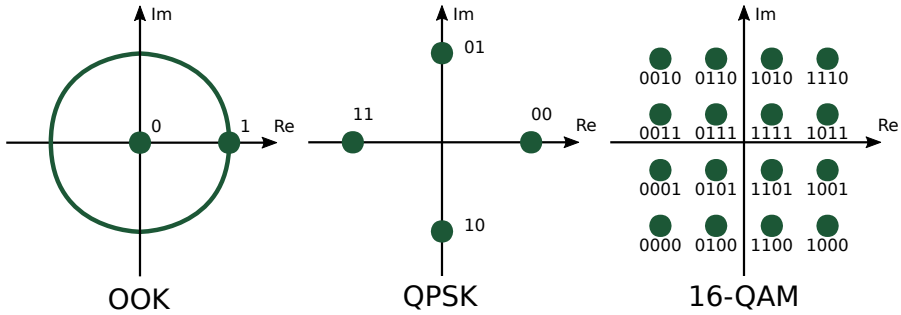


Figure 2.2: Constellation diagrams showing on-off keying, quaternary phase-shift keying, and 16-quadrature amplitude modulation. On-off keying is phase agnostic, while the other formats require phase recovery. Here, gray labelling is employed.

nents of the transmitted pulses, results in broadening of the transmitted symbols, causing the symbols to overlap and thus inter-symbol interference (ISI) which needs to be corrected. The resulting ISI causes a significant amount of symbols to overlap, in long-haul links, typically in the order of hundreds of symbols. Due to the attenuation, the signal requires amplification; however, these amplifiers add noise to the signal. The refractive index of the fiber is also power dependent, which in combination with chromatic dispersion causes distributed power-dependent phase rotation of the transmitted signal, and thus causing a nonlinear response with memory.

In addition to chromatic dispersion, the propagating wave is also affected by polarization-mode dispersion (PMD). The refractive index varies with polarization due to nonidealities in manufacturing and mechanical stress, which leads to an over-fiber polarization-varying dispersion and distributed polarization rotations, and thus pulse proadening. As this effect depends on external factors, such as temperature and vibrations, this effect needs to be compensated for using dynamic equalization. Additionally, since the effect causes interference both in time and over polarizations, both polarizations needs to be taken into account when compensating for PMD. Typically, PMD changes rather slowly [18]; however, there may be sudden rather rapid changes in state-of-polarization due to mechanical stress [19] or lightning strikes [20]. There are several approaches to modeling PMD and state-of-polarization rotation effects (for example, [21, 22]); however, sparsely occurring events which cause rapid changes may affect performance significantly. Instead, it may in many cases more useful to evaluate compensating algorithms by evaluating the tracking of a deterministic rotation, as this approach yields a clear comparable algorithm performance metric.

As earlier mentioned, fiber-optic communication suffers from noise. Typically, coherent transmission systems are limited by the added amplified spontaneous emission noise in amplifiers [23, Ch. 16], whereas short-reach IM/DD

systems are commonly limited by thermal noise [23, Ch. 4]. In long-haul applications, lumped erbium-doped fiber amplifiers (EDFA) are commonly employed with amplifiers distributed along the link. Thus, when nonlinearities are taken into account, noise needs to be added at each simulated amplification point as nonlinear signal-noise interaction is important. On the other hand, if only linear single-channel effects are taken into account, noise can be added to a single point in the simulated fiber. The quantum limit regarding EDFAs is a noise figure of 3 dB [23, Ch. 7]; commercial units with a noise figure of <5 dB are available [24, 25]. In addition, in the case of dispersion unmanaged links (where the bulk CD compensation is performed digitally) , uncompensated nonlinearities may be modeled as additive Gaussian noise [11, 26, 27]. However, since some of the nonlinear effects are deterministic, compensation of nonlinearities is possible and can yield better system performance. Another important noise-like effect, which is not due to the fiber-optic channel itself but rather analog-to-digital (ADC) and digital-to-analog converters (DAC), and DSP, is quantization of the signal and numerical rounding in the implemented algorithms. It is important to discern between signal rounding and rounding of static coefficients, as the latter leads to static errors (for example, filter-coefficient rounding causes deviation in filter response). In case of signal rounding, the effects are clearly signal-dependent and deterministic; however, it is in many cases useful to treat quantization as random noise.

In modern systems, free-running transmitter and receiver lasers are employed, which is referred to as intradyne systems. The lasers suffer from phase noise and frequency drifts which requires compensation in the receiver. Commonly, the compensation is split into two stages: frequency offset compensation and phase-noise compensation. For example, in [28], coarse carrier recovery is performed early on in the chain, and fine carrier recovery is performed later on to handle more rapid fluctuates. The frequency offset of the lasers cause an offset of the digitized spectrum, and be modeled as a static offset. Phase-noise is typically modeled as a Brownian walk of the phase in the signal [29].

If we disregard FEC for a while, increasing reach requires increasing the SNR and thus either reducing noise or increasing input power. However, the fiber is nonlinear, and increased launch power increases nonlinear impairments. Nonlinear signal-signal interaction is deterministic and can be compensated for, given that we know the received interacting signal. In practice, receiver bandwidth is rather limited (state-of-the-art A/D conversion can achieve around 90 Gsamp/s [30]), and we are thus limited to compensating in-band interactions. Even in this case, effective algorithms are quite complicated and even if we assume full knowledge of the entire WDM spectrum and perfect nonlinearity compensation, we are eventually limited by nonlinear signal-noise interaction [31]. Thus, in order to increase reach, we need to be able to correct bit errors at the output using forward error correction. Modern forward error correction schemes can operate fractions of dB from the Shannon limit; however, even relatively weak codes such as the Reed-Solomon codes used in older generation long-haul systems may improve performance significantly. In modern

long-haul coherent systems, it is common to use high coding gain soft-decision systems based on low-density parity check (LDPC) codes or turbo-product codes. For shorter reach systems, hard-decision decoded product or staircase codes are commonly considered.

2.1.2 Employed System Models

Since the focus of this thesis is energy-efficient DSP and FEC algorithms and implementations for fiber-optic communication systems, the goal is low-power, good compensation and correction rather than striving for the best possible performance. Thus, as earlier mentioned, rather than having advanced models which captures all possible effects and impairments in realistic systems, we want models that present an (from the algorithm perspective) idealized case and focus on the relative performance and power dissipation of the algorithms.

In **Paper A**, we focus on the effect of non-ideal filter implementation (finite length effects and rounding due to limited word length) and we employ a dispersive AWGN channel to estimate performance loss to the ideal case. The model is also employed to generate test data with proper statistics for power simulation of the circuit implementation. **Paper B** focuses on implementation of dynamic equalization, and we here instead use a linear AWGN channel with deterministically rotating state-of-polarization of a varied angular velocity, which allows us to quantify the performance of the parallel, pipelined dynamic equalizer implementation as well as the performance reduction due to computational simplification.

Paper C–F focuses on nonlinear mitigation. Here, a single-channel, single-polarization nonlinear split-step Fourier method model with lumped amplification are employed. The model presents a best-case scenario for nonlinear mitigation algorithms, and allows us to quantify the effect of both signal and coefficient rounding errors as well as arithmetic simplification. Additionally, in the papers concerning circuit implementation, the models are also employed to generate power simulation test vector with proper switching statistics.

In contrast to the other papers, **Paper G** focuses on simple IM/DD VCSEL-based links, which employ hard-decision slicers and are commonly limited by thermal noise. The channel is thus assumed to behave as a BSC channel. In addition, since errors are very sparse at the assumed power levels, the power dissipation is evaluated in an error-free channel with uniformly distributed encoded data. **Paper H** also considers hard-decision decoding, and thus a BSC channel; however, here the bit-error rate is much higher and has a significant impact on decoder utilization and therefore also power, and is thus included in the evaluations. **Paper I** uses soft information to assist the decoders. Here, we employ a BI-AWGN channel. The output of the channel is quantized to data and single-bit reliability before being inputted to the decoder.

Finally, **Paper J** considers system-level aspects and employs a linear, dispersive channel with phase noise. All considered algorithms are instantiated in context and operates on the previous-algorithm processed data. For the FEC

considerations, hard-decision decoding with long interleaving is assumed, and FEC is thus evaluated separately at the estimated channel output BER. The model is also used to generate test data for power simulation of the included algorithms.

Chapter 3

Digital Signal Processing

While DSP can be implemented using many different technologies such as DSP processors, field-programmable gate arrays (FPGA), and application-specific integrated circuits (ASIC), at the high throughput requirements and energy and power limitations, only ASICs are feasible for practical systems; ASIC design is thus the underlying consideration of this thesis. Modern ASICs contain millions of transistors, and circuit implementation is clearly a daunting task. This chapter provides an introduction to CMOS integrated circuits, and semi-custom ASIC design, as well as algorithm implementation consideration. The focus is primarily on real-time ASIC implementation; however, it should be noted that the implementation considerations discussed here are relevant for other DSP implementation styles as well.

3.1 CMOS Integrated Circuits

Complementary metal-oxide semiconductor (CMOS) integrated circuits is a key enabler of very large scale integration (VLSI) circuits. CMOS provides dense integration and by virtue of to the complimentary pair operation, the possibility of very low power dissipation; CMOS has therefore become the mainstay technology in digital integrated circuit design. An essential benefit is that the complimentary operation provides no direct path from the supply rail to ground. Thus, when signals are static, the only current that flows is due to leakage. In addition, each logic gate provides a full-swing output.

The power dissipation of CMOS circuits consists of three main components, dynamic, static, and short-circuit power dissipation. Typically, dynamic power dissipation tends to dominate overall power dissipation in stream-processing circuits (such as DSP), while static power dissipation is a concern in circuits where processing is performed relatively sparsely (such as FEC). Short-circuit power can typically be disregarded in deep submicron CMOS logic circuits [32, Ch. 8].

Dynamic power dissipation is caused by charging and discharging of the transistor gates in logic circuits, and can be expressed as

$$P_{sw} = fC_{\alpha}{V_{DD}}^2, \tag{3.1}$$

where f is the clock rate of the circuit, C_{α} is the sum of all circuit capacitances times their respective switching probability, and ${V_{DD}}^2$, is the supply voltage of the circuit. C_{α} depend both on the circuit complexity (more transistors, and thus capacitance) and speed requirements (since higher speed requires higher drive strength, capacitance is higher) and on signal statistics, as power is dissipated when switching occurs.

Static, or leakage, power is dissipated regardless of signal activity. Thus, static power is a major concern in circuits where large parts may remain relatively inactive, such as memories or FEC decoder back-ends. The main design parameters regarding leakage power is supply voltage and device threshold voltage; however, both parameters also significantly affect device speed, and careful design and consideration of the speed-power trade-off is thus required. Typically, for the technologies considered in this thesis, it is desirable to use higher-threshold devices for FEC circuitry. In contrast, in streaming DSP, due to the large switching activity in streaming DSP, lower threshold devices are beneficial as the higher speed can allow for use of smaller devices with less gate capacitance.

3.1.1 Semi-custom ASIC design

Modern integrated circuits can contain millions of transistors; obviously, full-custom design—drawing all features manually—is out of the question. Instead modern digital ASICs employ a semi-custom design flow where the full digital circuit is implemented using small standardized, foundry provided, logic cells with layouts, so called standard cells. Instead of starting with a schematic and drawing a corresponding layout, semi-custom design starts with a description of function and connections in a hardware definition language (HDL), where the two most common are VHDL and Verilog (which is nowadays a subset of SystemVerilog). Both VHDL and Verilog are dataflow languages, similar to circuit netlist descriptions with added behavioral description of parts. The HDL is then synthesized to a gate-level netlist using heuristic algorithms for logic optimization and gate sizing. The resulting gate netlist is then place-and-routed into a layout. Fig. 3.1 shows a place-and-routed ASIC implementation of a product decoder, with a fully-custom padframe.

Modeling and estimation of performance can be performed at each step, with increasing accuracy as the flow is progressed. Early on, block level switching can be estimated, but there is no notion of actual circuit elements. After synthesis, a netlist with gates is available, along with statistically-estimated wire loads. While wire lengths may be estimated decently on average [33, 34], there are still discrepancies that may show up at the time of physical placement. After place-and-route, the layout is finished and the circuit parasitics may be

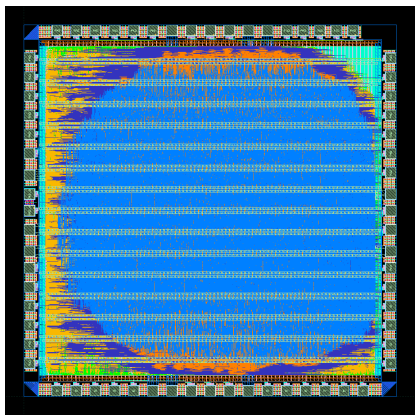


Figure 3.1: Example of a place-and-routed semi-custom ASIC layout in a custom padframe. The design contains around 10 million transistors.

extracted for accurate estimation of performance. In this thesis, estimations are performed at synthesis level, due to the manual tweaking and very long runtime that is required for full place-and-route. We have compared physical placement-aware estimations for small circuits, which indicated good correlation with statistical models; however, it should be noted that there can be some variations on large designs due to physical locations of processing elements and memories.

3.2 Digital Signal Processing

As earlier discussed, the fiber-optic channel suffers from several impairments, and effective mitigation of impairments is thus essential for reliable communication. Traditional fiber-optic communication systems relied on square-law detection, which does not allow for full reconstruction of the electromagnetical field envelope, and thus required optical compensation methods for compensating impairments such as chromatic dispersion. In contrast, coherent systems allow for linear capture and digitization of the full electromagnetical field envelope, and thus enables effective compensation of impairments in the digital domain and the use of high spectral efficiency quadrature-amplitude modulation (QAM).

While the fiber is a nonlinear channel, the systems are generally operated in a pseudo-linear regime and the DSP structure is typically designed from a linear perspective. Both linear impairments with memory such as chromatic dispersion and memoryless nonlinear impairments such as nonlinear responses in modulators may be compensated in a relatively straight-forward fashion, for example using linear filters and nonlinear predistorsion, respectively. However, real-time compensation of impairments with a significant nonlinear memory

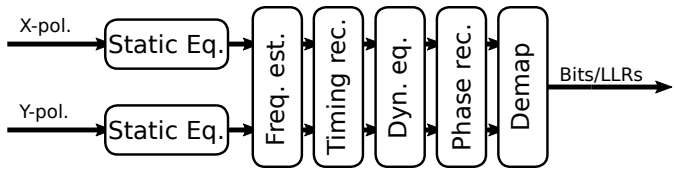


Figure 3.2: Simplified block diagram of a receiver DSP chain.

such as Kerr nonlinearity in combination with chromatic dispersion remain a significant challenge. Commonly in fiber optic communication research, DSP is performed offline on batches of data captured during experiments. In this case, throughput is rarely taken into account and many complex algorithms involving (in some cases iterative) symbol-by-symbol processing is usually applied. However, since this thesis focuses on energy efficiency aspects, we need to focus on current ASIC technology real-time realizable algorithms and implementations. Thus, we here focus on algorithms that are mainly feed-forward or block feedback. Fig. 3.2 shows a simplified DSP receiver structure. First, static compensation of chromatic dispersion (and in some cases nonlinear compensation) is performed. Afterwards, the frequency offset is compensated and sample timing is recovered. Then, dynamic equalization is performed to compensate for PMD and other residual impairments (additionally, the dynamic equalizer inherently also performs some sample-phase compensation if there is any offset). Signal-phase recovery is then performed and the received samples are demapped either to bits or, if soft-decision FEC is employed, log-likelihood ratios.

Chromatic dispersion acts as an all-pass filter with quadratic phase, and is corrected by convolving the signal with a filter that has the inverse response. At high symbol rates and long fibers, the impulse response of the required filters is long (in the order of hundreds of taps) and is typically compensated using overlap/save frequency-domain processing. Here, fast Fourier transforms are used to efficiently perform cyclical convolution on overlapping blocks of the data. Parts which are affected by artifacts are discarded, resulting in linear convolution. In terms of performed multiplications, overlap/save is very beneficial. However, when fixed-point math is employed, the application of the twiddle factors requires an increase in word length in order to keep the signal to noise ratio (SNR) reasonable [35], and a comparison based on multiplicative complexity may thus be rather misleading. Thus, for shorter links, time-domain approaches may be of interest.

While chromatic dispersion is straight forward to correct, the interaction of CD and fiber nonlininearity is not. The nonlinearities causes a distributed phase rotation that depends on the instantaneous power of the field, which is in turn affected by CD. Thus, the nonlinear response of the fiber causes nonlinear impairments with memory. In order to compensate for this, we can employ simulated propagation with negated fiber parameters, taking reasonably small

dispersive steps intertwined with nonlinear phase rotation, commonly referred to as digital back propagation (DBP) [36]. In practice, since the algorithm needs to operate in a streaming fashion, each step consists of a linear filter, which correct a small amount of dispersion, concatenated with a block performing a power-dependent rotation. The steps are cascaded to compensate for the full link. In theory, a smaller linear step, and thus more steps for the total link, leads to better compensation; however, in practical fixed-point applications, numerical effects will impact the over-all performance significantly. Other options include perturbation approaches [37], and compensation based on Volterra series [38]. However, while interesting, these approaches are not considered in this thesis.

Since the receiver and transmitter lasers as well as the sampling clocks are free running, there is an offset in frequency and sampling time that needs to be estimated and compensated for. Possible approaches to frequency-offset estimation include 4th-power estimation [39], FFT-based estimation [40], and coarse FFT-based compensation followed by a gradient descent algortithm [41]. Regarding timing recovery, classical methods include the Mueller and Muller method [42] and the Gardner method [43]. The actual sampling-time compensation can be performed using interpolation in the digital domain [44]; in this case, it needs to be ensured that the receiver clock runs faster than the transmitted clock in order to enable lossless interpolation. The compensation can also be performed mixed-mode, by controlling the ADC sampling clock using the timing-recovery algorithm [28].

Once timing recovery has been performed, we need to compensate remaining linear effects using dynamic equalization. While commonly considered to primarily target PMD and polarization demultiplexing, the dynamic equalizer serves as a catch-all compensator for remaining linear effects. The dynamic equalizer commonly consists of a 2×2 complex multiple-input multiple-output FIR filter, with taps continuously updated. The filter can be expressed as

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \mathbf{h}_{11} \, \mathbf{h}_{12} \\ \mathbf{h}_{21} \, \mathbf{h}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} \tag{3.2}$$

where $\mathbf{h}_{ij}$ are the continuously updated filter taps. It is also possible to employ a real-valued 4×4 filter; such structure has the benefit that it can compensate for skew between the in-phase and quadrature components of the received signal (IQ-skew) [45]. Tap update calculation can be performed either blind (commonly using classical algorithms such as the constant-modulus algorithm (CMA) [46], radius-directed equalization (RDE) [47], or the decision-directed least mean square algorithm (DD-LMS) [48]) or using pilot symbols [49]. CMA and RDE are phase-agnostic and only employ amplitude information, and the error can thus be calculated at the output of the equalizer. In contrast, DD-LMS requires proper decisions and thus requires phase compensation to be placed within the dynamic equalization tap-update feedback loop.

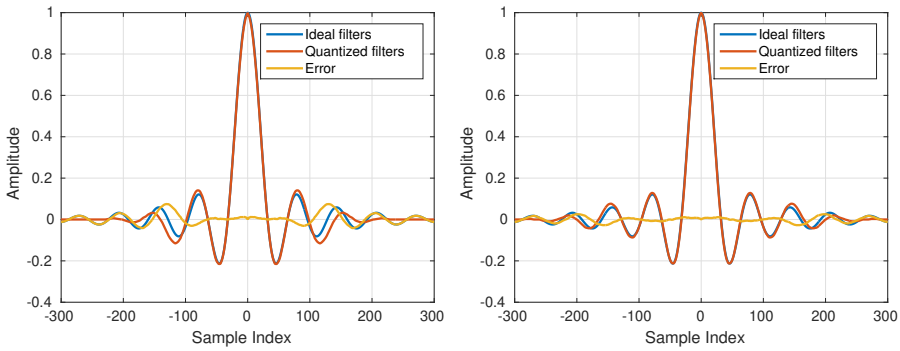
Once linear impairments are compensated for, signal-phase offset can be estimated and compensated for. Common approaches to phase-noise compen-

sation include the Viterbi-Viterbi algorithm [50], blind phase search [51], and pilot-based estimation [49]. While blind methods does not require insertion of known symbols (and thus reduction of spectral efficiency), they suffer from a finite probability of making catastrophic errors due to rotational symmetry of constellations, so-called phase slips. In contrast pilot-based schemes avoid this issue at the cost of sacrificing some spectral efficiency. In addition, pilot-based approaches are rather suitable for real-time parallel implementation [52]. Finally, the eye has been properly opened and the received symbols can be demapped either to bits or log-likelihood ratios.

3.2.1 DSP Implementation Aspects

Typical CMOS circuits operate at a clock rate in the order of 1 GHz, whereas optical communication systems requires processing of around 50 Gsamp/s. Extensive parallel processing and pipelining (cutting timing paths with registers to increase the maximum clock rate) is thus required. While algorithm design commonly considers high-resolution floating-point processing, practical real-time DSP requires implementation of algorithms in fixed-point arithmetic in order to limit circuit complexity and power dissipation. The circuit complexity of the required fixed-point arithmetic units depends heavily on the word lengths of the operands, and resolution requirements thus plays a significant role in the algorithm power dissipation. However, reducing word lengths increases rounding errors, thus giving rise to a performance-power trade-off. In DSP circuits, we need to differentiate between rounding of signals and rounding of static operands such as filter coefficients. Although both are systematic errors, the former is signal dependent and behaves noise-like (commonly referred to as quantization noise), the latter causes a systematic signal-independent error.

Consider a typical root-raised cosine pulse shaping filter. The convolution of the transmitter pulse-shaping filter and the receiver matched filter should should result in a over-all impulse response that fulfills the Nyquist criterion. However, when implemented in limited-resolution arithmetic, rounding errors cause impulse-response deviations causing the over-all pulse-shaping/matched-filter pair response to deviate from the ideal case. In the following example, we assume that the pulse-shaping and matched filter is implemented using 4-bit fixed-point arithmetic, oversampled to 32 SPS for figure clarity. Fig. 3.3a shows the over-all ideal pulse-shaping and matched filter response, the response with equally-quantized filters in both transmitter and receiver, and the corresponding error power. The total error power is approximately -4 dB. Equally-quantized filters are clearly suboptimal as the errors add coherently. If we instead first create a quantized pulse-shaping filter, we can then transform the filter into frequency domain to find a filter frequency response that cancels the induced quantization errors. We then transform the filter back to the time domain and quantize to the same resolution as before to obtain a different matched-filter approximation. Fig. 3.3b shows the over-all ideal pulse-shaping and matched filter response, the response with unequally-quantized filters in



(a) Impulse response with ideal filters, equal quantized filters, and the corresponding error.

(b) Impulse response with ideal filters, unequal quantized filters, and the corresponding error.

Figure 3.3: Over-all floating-point and 4-bit quantized pulse-shaping/matched filter impulse response.

the transmitter and receiver, and the corresponding error power. Here, we have reduced the over-all error power to approximately -10 dB. In the example here, we only consider two cascaded filters; in the case of algorithms containing a large cascade of operations such as DBP, co-designing quantized filters may yield a much larger benefit. However, rounding is a nonlinear operation, and co-optimizing a large amount of fixed-point filters is difficult at best.

As earlier mentioned, the high symbol rates of modern fiber-optic communication systems requires extensive parallel processing and pipelining in DSP algorithms. Broadly, we can classify processing as either feed-forward or feedback systems. Feed-forward systems are preferable in high-throughput applications, as they can be parallelized and pipelined fairly straightforwardly with throughput mainly limited by silicon area and power. Fig. 3.4 shows an example of a parallelized FIR filter structure. In contrast, feedback loops puts a strict upper bound on the achievable throughput [53]. Thus, algorithms that require feedback loops, such as dynamic equalizers, needs to be modified to reach the required throughputs. For example, in a typical LMS-style equalizer, coefficient update is performed after each filtered sample, thus inferring a limiting feedback loop. In order to implement the algorithm, we can instead modify the algorithm to only update coefficients block-wise, thus avoiding this issue. Block processing obviously limits the update rate (and thus tracking speed) somewhat, but in practice, this is not an issue at typically considered signal-to-noise ratios, as the main limitation regarding tracking speed is the signal noise content.

Hardware implementation of transcendental functions poses a significant complexity-accuracy trade-off. The functions are commonly implemented using look-up tables or polynomial expansion approximation (in some cases, with interpolation to improve accuracy) [54], or using the CORDIC algorithm [55].

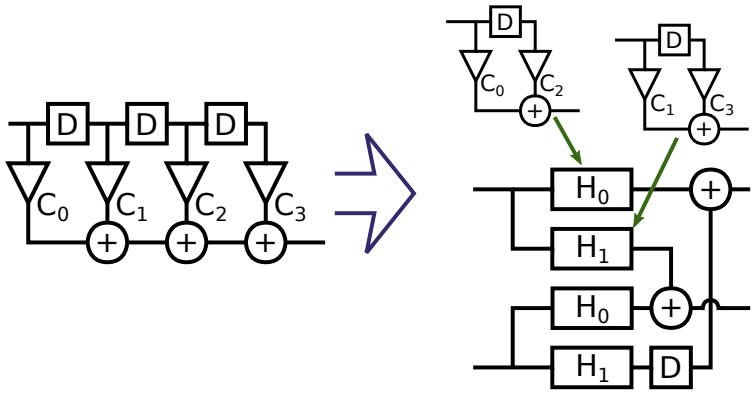


Figure 3.4: Block diagram illustrating parallelization of a four-tap FIR filter. The filter is decomposed into even and odd coefficients, and two sets of the resulting two-tap filters are combined into a two-parallel structure.

It is also possible to employ logarithmic arithmetic units [56]. The resolution and numerical accuracy will clearly affect the size of look-up tables or the required number of terms in the polynomial expansion. In this thesis, complex exponentials are required in DBP. Fortunately, the required rotation in each step is rather small, and can be implemented with few-term Taylor expansions.

3.3 Forward Error Correction

Whereas systematic impairments can be compensated for, random noise remains. Forward error correction (or error control coding) allows for the detection and correction of errors, up to a certain error threshold, by introducing redundancy in the transmitted data stream. Richard Hamming invented the first forward error correction code, the Hamming code, in 1950 to prevent computers from stopping calculations when errors were detected [57]. Hamming arranged three parity bits in such a way that when a single error occurs in a seven-bit block, the position of the bit flip can be found by checking which parity constraints that are fulfilled. Since Hamming's pioneering work, many codes have been invented. In fiber-optic communication, the most commonly considered algebraic codes are BCH codes [58, 59], operating on bits, and Reed-Solomon codes [60], operating on symbols. The codes are commonly used as component codes for construction of longer block-length codes with reasonable decoding complexity such as product codes [61] or concatenated codes [62]. In the early 1990's, Turbo codes [63] revolutionized coding theory and presented a paradigm shift; in contrast to algorithmic coding where algebraic structures of the code was exploited for decoding, Turbo codes iteratively processed likelihood ratios. Turbo codes sparked a great interest and lead to the rediscovery

of LDPC codes, invented by Gallager in the 1960's [64]. Remarkably, Gallager presented iterative belief-propagation decoding in his thesis; however, in 1963, integrated circuits were in their infancy and the algorithms were thus too complex to use, and the codes were for a long time forgotten. Nowadays Turbo and LDPC codes, and iterative belief-propagation decoding have become a mainstay in modern communication systems.

Traditionally, codes were classified either as convolutional (with undefined length) or as block code (operating on a fixed-size block). Compare, for example, BCH codes and convolutional codes, or in the case of modern codes Turbo and LDPC codes. In fiber optics, block codes are typically employed as the block-wise operation allows for fast, parallel encoding and decoding. However, nowadays several popular codes such as staircase codes and spatially-coupled LDPC codes combine features from both code classes. Staircase codes, which are considered in this thesis, consists of chained overlapping blocks, and are decoded using block-wise decoding operating on a small section of the long chain, commonly referred to as windowed decoding.

Decoding algorithms are typically categorized as hard-decision, in which the decoding operates on a quantized bitstream, or soft decision, where decoding operates on reliability metrics. A third option that has recently received attention is the possibility of employing hard-decision algorithms as a core, and then assist the decoding using some soft information, which we refer to as soft-assisted decoding. Typically, soft-decision algorithms are rather complex in comparison to hard-decision decoding, and hard-decision decoding may thus be implemented more power efficiently. However, soft information is inherently available in DSP-based coherent receivers; soft-assisted algorithms utilize the available information to assist a core hard-decision algorithm using low-complexity hardware.

3.3.1 Product-like codes

Product codes, invented by Peter Elias in the 1950's [61], offers a way to practically build long powerful codes using simpler short component codes, and offers high coding gain when decoded using iterative hard-decision decoding. Fig. 3.5 illustrates the encoding process of a product code using classical Hamming codes as component codes. Data to be transmitted is placed in a square array (marked in green) and each individual row is then encoded using the component code, generating row parity, the array is enlarged in the row direction, and the parity is stored (marked in blue). All the columns of the resulting array (the data and row parity) is then encoded column-wise using the component codes, the array is then extended in the column direction and the resulting column parity is stored (marked in red). This structure results in an overall code with a minimal distance of $d_{min} = d_{min_1} \cdot d_{min_2}$, where d_{min_1} and d_{min_2} are the minimum distances of the row and column codes, respectively. Commonly, product codes employ the same code for both row and column component codes.

When the data is received, the product decoder iteratively decodes the

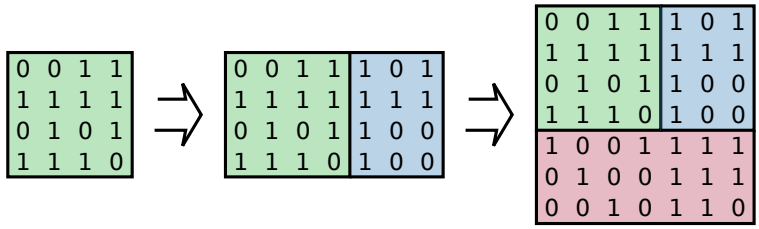


Figure 3.5: Illustration of the product code encoding process. The information bits are marked in green, row parity in blue, and column parity in red. First, all rows are processed separately to generate the row parity. Then, all columns are processed separately to generate column parity.

component codes. Each iteration consists of first decoding all rows (including parity-on-parity) separately using a decoder for the employed component codes. Found and correctable errors are then corrected accordingly. Once all rows have been decoded, the same procedure is performed for all column codes separately. This process is then repeated until all errors are corrected or a set maximum of iterations are reached. Fig. 3.6 illustrates the iterative decoding process of the received Hamming-code-based product code block. First, all rows are decoded. Since the Hamming code can only correct one bit error, the first iteration can only correct one error in the particular error pattern in this example; the row component code affected by two errors is uncorrectable. In the second half-iteration, all columns are decoded. In the example here, the remaining bit errors belong to separate component codes and can subsequently both be corrected by the column component decoders. Here, one iteration (consisting of a row half-iteration and a column half-iteration) is enough to correct all errors; however, typically up to 5–10 iterations per received block are allowed. In this example, all errors were correctable; although, if instead a 2×2 rectangular pattern would be received (for example, if the top-left bit was also received erroneously), decoding would not be able to correct the errors. The considered product code has a minimum distance of $3^2 = 9$ and should thus be able to correct 4 errors. This is commonly referred to as a stall pattern, and causes error floors. In practical systems, component codes capable of correcting multiple errors are employed in order to ensure that the error floor is below 10^{-15}. While iterative decoding cannot decode the over-all product code up to its maximum error-correcting capability, it offers a hardware-friendly way of creating well-performing practical error-correcting schemes.

While the performance of these codes is typically very good, further increase of performance is possible by introducing spatial coupling of blocks. A relatively recent development are Staircase codes [9], which places the bits in an overlapping staircase-fashion, forming a product-like code with increased performance while still being decoded using simple hard-decision algorithms. The first sub-block is filled with all zeros, and encoding commences in a chain from

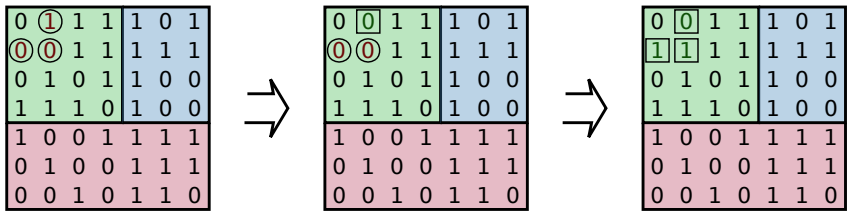


Figure 3.6: Illustration of the product code decoding process. Received bit errors are marked in red and circled, and corrected bits are marked in green and squared. Note that the bottom-left corner of the column parity corresponds to parity-on-parity and is the same regardless of whether row-first or column first encoding is performed.

this point. Due to the shortening in the beginning of the chain, the component codes are effectively stronger at this point; the shortened part acts as a catalyst for decoding, starting a decoding wave that is propagated along the chain. To decode staircase code, we employ windowed decoding. The decoder operates over a small part of the chain, the decoder window. A part of the chain of overlapping blocks are loaded into the decoder window, and the in-window blocks are iteratively decoded. After iteratively decoding the in-window blocks, the window is shifted one step and the oldest block is shifted out while a newly received block is shifted into the window. As earlier mentioned the first block in the chain consists of all-zeros (this part is not transmitted, the block is thus shortened, and no errors can occur here); thus, there is a higher probability of decoding success in the second block whose component codes overlap with the first. At this starting point, the lower bit-error boosts the decoding process of the first blocks. Due to the spatial coupling, a decoding wave forms, reducing the BER in the blocks in the end of the decoding window, thus boosting the decoding of the following blocks. Fig. 3.7 shows an illustration of the staircase code structure.

3.3.2 FEC Implementation Aspects

In contrast to typical stream-processing DSP algorithms, FEC commonly employs iterative processing, where processing units are only activated if necessary. Thus, both dynamic and static power is of concern. As CMOS power dissipation is highly dependent on the switching activity, it is essential to minimize data movement in order to reduce power dissipation. In addition, clock gating is highly beneficial; the large clock trees may be a significant contributor to overall dynamic power, and due to the error-detection/correction structure of algorithms, there is ample opportunities in early detection of logic that can be disabled.

Since the throughput requirements of fiber-optic communication systems are immense, careful consideration of algorithm design is necessary. Since the

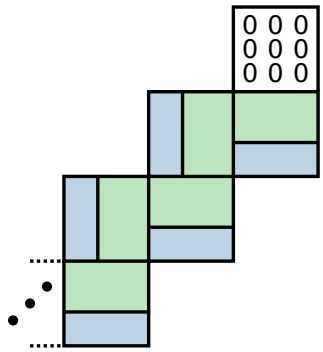


Figure 3.7: Illustration of the staircase code structure. As with product codes, decoding is performed in a row-column fashion.

algorithms rely on iterative processing, we can process the data in place and only flip corrected bits, or we can unroll the iterations and implement the design in a more streaming fashion. In the case of in-place processing, the throughput is bounded by the number of iterations and the processing time of one iteration; very fast component decoders are thus required in order to reach high throughputs. In iteration-unrolled designs, each consecutive iteration has its own hardware, and the output data from one iteration is moved to the next. Here, high speeds are possible at the cost of replicated hardware for each iteration. In addition, comparing the two approaches, it is clear that an iteration-unrolled design is likely to be less power efficient than an in-place processing design: since the data is not correlated over blocks, the required data movement causes high switching activity and may thus lead to excessive power dissipation. In this thesis, in-place architectures with fully-parallel component decoders are considered. These architectures are very power efficient, at the cost of rather high area requirements due to the many component decoders employed. The presented architectures are mainly suitable for moderate-to-high overhead codes (20% or higher). While not considered in this thesis, sharing decoder back-ends between independent rows and columns could yield a large area reduction since back-ends are fairly large; however, the required control logic will be more complex.

Comparing hard-decision algorithms to soft-decision algorithms, it becomes apparent that not only are HD algorithms less complex, since only detected errors needs to be flipped in the case of in-place processing, HD algorithms typically lends themselves well for implementation. In contrast, soft-decision algorithms require update of all in-code likelihood ratios (which are commonly encoded as fixed-point numbers) and thus leads to significant switching activity.

Chapter 4

Fiber-Optic Communication Sub-systems

Since the early days of fiber optic communication, there has been a rapid increase in system throughput thanks to inventions such as the low-loss fiber [2], erbium-doped fiber amplifiers [65], and relatively recently, DSP-based coherent systems [3]. Commonly, systems were classified as short-haul systems, mainly employing simple direct-detect schemes, or long-haul, employing complex modulation format and impairment compensation; however, advanced technologies are nowadays considered for shorter and shorter reach systems. Still, shorter reach systems require less chromatic dispersion compensation and simpler dynamic equalization, and DSP can be tailored to the specific application in order to reduce power dissipation. In addition, simpler lower-performing hard-decision decoding forward error correction can be employed as well, further reducing receiver power dissipation.

Modern coherent fiber-optic communication systems consists of several complex subsystems, ranging from optical and electro-optical devices such as fibers, optical hybrids, modulators, and EDFAs, to electrical integrated circuits and mechanical packaging and cooling. In the early days of digital coherent systems, transceivers consisted of a large amount of components, including several ASICs, integrated on a printed circuit board [66]. Since then, there has been a remarkable progress regarding integration and miniaturization of both optical and digital subsystems; nowadays, coherent transceivers are available in small packages such as CFP2 [67]. Still, further integration and power dissipation reduction is essential in order to fit coherent systems into even smaller packages such as QSFP-DD [68] or OSFP [69]. Heat management in such small packaging is difficult, and power dissipation in digital subsystems, and in particular the receiver, is therefore an important concern.

The systems are complex, and there are clearly many important sub-systems other than the systems considered in this thesis. The output from the DSP is a bitstream (or stream of LLRs), and buffering and framing needs to be performed as well. In addition, electrical interfaces between the receiver ASIC and the host is required as well, including buffering, serialization, deserialization, and synchronization. High-speed design of these short electrical links is a significant effort in itself, and any effects due to these links are not considered here.

4.1 System Power Dissipation

DSP and FEC contributes to a significant amount of power dissipation in coherent fiber-optic communication transceivers [70, 71]. In the case of long-haul coherent systems, DSP and FEC has been estimated to account for up to around 50% of the over-all power dissipation of the system [1]. CD compensation, dynamic equalization, and FEC are considered to be the most complex blocks in the receiver ASIC [72], and have been estimated to contribute to a majority of the receiver power dissipation [1, 73]. In the case of the ASIC receiver implementation presented in [28] (which does not include FEC), CD compensation dissipates 44% of the overall power, while the dynamic equalizer dissipates 28%. This thesis thus mainly focuses on CD compensation, dynamic equalization and FEC. While not considered in the estimations, nonlinear compensation, which is even more complex than CD compensation is also considered. In order to isolate implementation effects of single blocks, this thesis mainly focuses on these sub-blocks separately (see **Paper A**–**I**); however, several blocks have also been integrated in order to estimate the feasibility of DSP-based coherent datacenter interconnects (see **Paper J**). It should also be noted that savings in sub-blocks also reduces the need for cooling and losses due to inefficiencies in power supply circuitry; however, such related savings are not taken into account here. Other major sources of power dissipation include the laser and modulators [1], and the link amplifiers [1, 74].

Since the systems require rather complex hardware and dissipate a significant amount of power, coherent communication have mainly been considered for long-haul systems. Despite this, recent trends indicate that coherent systems will reach into lower-distance links such as datacenter interconnects [4, 75, 76], links which are densely packed and thus power constrained. Since DSP contribute to a significant amount of power dissipation, DSP-free coherent systems have sparked interest recently, both for intra-datacenter links [77, 78] and inter-datacenter links [77]. However, DSP can provide flexibility, and as will be shown in **Paper J**, energy-efficient design and implementation of algorithms may enable the implementation of a full DSP-based coherent receiver within small power-constrained packaging.

4.2 Considered Systems and Algorithms

Paper A considers implementation of chromatic dispersion compensation using both time-domain and frequency-domain approaches. Here, word-length requirements are determined and parallel fixed-point filters which meet the throughput requirements are implemented. Power dissipation and area requirements are evaluated at the synthesis stage using statistic models, using data generated with the considered system model. **Paper B** considers dynamic equalization, using a similar synthesis-based approach. In contrast to the strictly feed-forward filters employed for CD compensation, parallelization and pipelining has a significant impact on the overall dynamic equalizer performance algorithm performance. Thus, in order to eliminate the risk of modeling discrepancies, the implemented VHDL equalizer is directly evaluated using a MATLAB/VHDL co-simulation approach.

Paper C considers the use of limited-precision arithmetic and function approximation in the implementation of digital backpropagation. **Paper D** considers the design of fixed-point filter pairs that partially cancels errors, and thus can achieve higher performance at lower resolution. Circuit implementation aspects of digital backpropagation are further explored in **Paper E**, where pair-wise optimized filters were considered, and in **Paper F**, where algorithms from a machine-learning framework were used to optimize the filters. Regarding ASIC implementation, both papers use a similar approach as employed in paper A.

Paper G considers fully-parallel high-throughput BCH-based FEC units in short-range vertical-cavity surface emitting-laser (VCSEL)-based links. Here, the circuits are small enough to employ a placement-based wire estimation flow. The decoders were further developed to be used as component decoders in high coding-gain hard-decision **Paper H** and soft-assisted **Paper I** decoders. These decoders were carefully designed to reduce data movement and unnecessary signal switching, and employ extensive clock gating. The considered staircase and product decoders are much larger and complicated as they consist of several component decoders, data memory, and control logic; here, due to run-time concerns, technology-data-based models in the synthesis tool are used.

Chapter 5

Contributions

5.1 Problem Statement

The over-arching concern of this thesis can be summarized as follows: How do we design and implement DSP and FEC algorithms and architectures that not only provide good compensation and correction performance, but also energy-efficient (or power-efficient) operation at the high throughput that is required by fiber-optic communication systems?

Since future technology scaling no longer promises to allow for implementation of increasingly complex algorithms, or conversely further reduction in power dissipation for current algorithms, it is now essential to explore efficient design of algorithms while bearing in mind resource limitations. Thus, in contrast to focusing on only pushing performance limits in terms of reach or bit-error rate or similar, we here focus on attempting to push several boundaries and explore trade-offs. In addition, high-level metrics such as arithmetic complexity comparisons may be misleading as neither implementation structure nor varying signal statistics are taken into account. Thus, circuit designs are investigated using modern CMOS technologies.

This thesis focuses both on the currently considered major contributors to receiver ASIC power dissipation as well as implementation of nonlinearity mitigation algorithms. Both high-level aspects such as resolution requirements and trade-offs, as well as lower-level aspects such as algorithm architectures and ASIC implementation, are considered. Algorithm design and implementation for energy efficient operation is investigated, and possible energy- and power-dissipation savings that can be expected if other performance constraints are relaxed.

5.2 Summary of Contributions

Paper A explores implementation of chromatic dispersion compensation for low-to-moderate amounts of dispersion using parallel FIR filters, fast-FIR filters, and frequency-domain overlap-save methods. Different filter design approaches are compared and finite-precision aspects are investigated. Filters are implemented and synthesized. Complexity-based metrics are found to not be suitable for comparing the considered implementations. It is shown that time-domain methods are mainly suitable for relatively short transmission distances, corresponding approximately to the step-sizes that are considered in digital backpropagation algorithms.

Dynamic equalization is explored in **Paper B**. Here, a pipelined, parallel-processing, dynamic equalizer, based on the CMA algorithm, is implemented. Power dissipation and tracking speed is investigated. It is found that even though the tap update algorithm and the filter has similar complexity, the tap-update block dissipates approximately twice the power compared to the filtering. In order to reduce power, we can use a reduced set of samples for tap-update calculation and remove the corresponding hardware, which we refer to as *sample pruning*. It is shown that sample pruning can significantly reduce over-all power dissipation, while still allowing the equalizer to track rapid changes.

In **Paper C**, design of nonlinearity mitigation suitable for finite-precision implementation is considered. Based on the knowledge obtained in paper A, time-domain optimized digital backpropagation is designed and finite-precision arithmetic aspects such as resolution requirements and function approximation is investigated. Here, instead of performing iterated Fourier transforms, all steps are performed in time domain. It is shown that time-domain digital back-propagation (TD-DBP) can achieve good compensation performance at rather moderate resolution. In general, DBP consists of many iterated steps, and rounding errors are therefore exacerbated. Applying pre-quantization dithering on the filter coefficients improves performance, showing that correlating errors are indeed a limiting factor. The implementation of the nonlinear operator is investigated, and it is shown that a simple first-order Taylor approximation provides good performance.

Since rounding errors are deterministic, they can be taken into account in the design of the fixed-point steps. In **Paper D**, co-optimization of filter pairs in the cascaded steps is considered. A novel metric, signal-to-interference ratio (SIR), is introduced and shown to correlate well with over-all system performance. Using SIR and a fast search algorithm, it is possible design quantized filter pairs that significantly outperform the single quantized-filter case. **Paper E** considers ASIC implementation of fixed-point optimized TD-DBP. It is shown that TD-DBP is feasible to implement, with similar energy dissipation as previously published estimations on CD compensation. While paper E shows that pair-wise optimized TD-DBP can be implemented rather efficiently, it is clear that there might be further gains by optimizing all cascaded steps in the TD-DBP chain. **Paper F** investigates ASIC implementation of

TD-DBP with steps optimized using algorithms employed in machine learning. Compared to pair-wise designed filters, similar performance is obtained using lower word-length filters with significantly fewer taps.

Paper G investigates possible system energy-dissipation reduction by employing simple BCH-based FEC in short-range VCSEL-based links that employ energy-efficient CMOS laser drivers. The added encoder and decoder dissipates very little energy, and it is shown that FEC can reduce over-all system power dissipation, at very modest decoding latencies. While this paper considers simple codes in short-distance links, the developed decoder architecture is suitable for use as component decoders in more advanced schemes. In **Paper H**, the decoders are further developed for 3- and 4-error correcting BCH codes, and both product and staircase decoders are developed and implemented. By carefully designing the architectures to reduce unnecessary data movement and signal switching, we implement very energy efficient decoders capable of >1 Tb/s throughput. In the case of the staircase decoders, it is shown that the majority of the dissipated power is due to the component decoders in the part of the window closest to the channel. In **Paper I**, the product decoders are further enhanced by using a small amount of soft information to assist the hard-decision core, which we refer to as *soft-assisted decoders*. By adding very simple logic, performance of the product decoders can be improved to reach coding gains similar to the more complex staircase schemes, in systems where soft information is available.

Finally, energy-efficient implementation of DSP for coherent DCI links are investigated in **Paper J**. A high symbol rate of 60 Gbd is employed, and non-integer sampling is considered to reduce power dissipation. Here, the dynamic equalizer from paper C is further developed for use with PM-16-QAM modulation, and integrated in the system model, along with frequency-domain CD compensation from [79], interpolation filters, and phase-noise compensation from [52]. The DSP implementations, product-code-, and staircase-code-based FEC are synthesized using a 22 nm process, and both power dissipation and area requirements are evaluated. It is shown that, in the considered system, FEC contributes to little of the over-all power dissipation. However, FEC requires a large part of the over-all silicon area. Regarding DSP, since the link length is short and the receiver operates at low oversampling rates, CD is relatively moderate and compensation can be performed rather energy efficiently; the major cause of power dissipation in the considered system is the dynamic equalizer.

To summarize, this thesis focuses on two major themes: implementation of DSP, and implementation of FEC. While both are parts of the same receiver ASIC, the algorithms are rather different regarding implementation concerns. The blocks commonly considered to contribute to the majority of the receiver ASIC power dissipation are investigated separately, and is shown that, using an implementation-focused approach, significant power-dissipation reduction is possible. Finally, both themes are put into a system context, and evaluated, and

it is shown that coherent DSP-based systems are feasible in power-constrained settings.

5.3 Future Outlook

While TD-DBP is suitable for fixed-point implementation, other nonlinearity mitigation algorithms such as perturbation approaches [37], and Volterra series [38] compensation have different algorithmic structures and are likely affected by rounding errors differently. It would thus be interesting to compare the different approaches in regards to numerical resolution aspects and circuit implementation. While these methods are effective for compensating intra-channel effects, inter-WDM impairments remains an issue. It has been shown that inter-channel nonlinearities manifests itself as a time-varying inter-symbol interference [80], and can be compensated for with fast equalization [81]. However, the algorithms considered in [81] are computationally complex and iterative; thus, it is likely that real-time implementation of such equalizers is difficult at best. Therefore, it would be interesting to investigate whether simpler parallel-processing equalizers can reach the tracking speed requirements.

The presented product and staircase decoders are very energy-efficient and can achieve very high throughput, but they are mainly suitable for moderate-to-high overhead codes, largely due to the many component decoders employed. It would thus be interesting to investigate further sharing of component decoders between separate component codes, and how to reduce switching activity in this case. Additionally in the case of staircase decoders, the decoders placed at the back of the window are sparsely used. It would thus be interesting to investigate if fewer decoders can be employed here. In this case, clever scheduling and multiplexing is required in order to prevent any reduction in correction performance due to temporary resource starvation, and power dissipation due to increases in switching activity.

Use of a limited amount of soft information in order to assist hard-decision decoders can be implemented with little added circuitry. However, the soft-assisted decoders considered here are still affected by the same stall patterns as the hard-decision decoders, and soft information is only used to reduce miscorrections. Further research on efficient use of soft information with limited added circuitry, not only for miscorrection prevention but also improvement of correction performance, is thus interesting. Recently, other relatively simple soft-decision and soft-assisted algorithms has been published [82–84]. However, these algorithms require sorting of received LLRs. Sorting or selection algorithms require many numerical comparisons and the cost of this added circuitry is still unclear. It would thus be interesting to investigate the cost, both in terms of power and throughput, related to sorting in such algorithms.

Large over-all power dissipation reduction of the dynamic equalizer can be achieved by simplifying the tap-update algorithm, but the filtering dissipates still a significant amount of power, especially if considered in the context of

DCI links. Compared to static filter implementation, filter implementation in dynamic equalizers not only needs to take power dissipation into account, but also processing latency as it is placed inside a feedback loop. Thus, frequency-domain or fast-FIR methods might not be suitable. Instead, since the tap update is fairly slow, it might be interesting to investigate the use of other number representations such as, for example, canonical signed digits, in the filter multipliers. While the conversion requires some added circuit, the impact of conversion of power dissipation might not be significant considering the slow changing taps, and thus low switching activity.

References

[1] B. S. G. Pillai, B. Sedighi, K. Guan, N. P. Anthapadmanabhan, W. Shieh, K. J. Hinton, and R. S. Tucker, "End-to-end energy modeling and analysis of long-haul coherent transmission systems," *IEEE J. Lightw. Technol.*, vol. 32, no. 18, pp. 3093–3111, Sept 2014.

[2] F. P. Kapron, D. B. Keck, and R. D. Maurer, "Radiation losses in glass optical waveguides," *Applied Physics Letters*, vol. 17, no. 10, pp. 423–425, 1970.

[3] M. G. Taylor, "Coherent detection method using DSP for demodulation of signal and subsequent equalization of propagation impairments," *IEEE Photon. Technol. Lett.*, vol. 16, no. 2, pp. 674–676, Feb 2004.

[4] M. H. Eiselt, A. Dochhan, and J.-P. Elbers, "Data center interconnects at 400G and beyond," in *Opto-Electronics and Commun. Conf. (OECC)*, Jeju, Korea, July 2018, pp. 6B2–2.

[5] E. Maniloff, S. Gareau, and M. Moyer, "400G and beyond: Coherent evolution to high-capacity inter data center links," in *Opt. Fiber Commun. Conf. (OFC)*, San Diego, CA, USA, 2019, p. M3H.4.

[6] R. S. Williams, "What's next?[the end of Moore's law]," *Computing in Science & Engineering*, vol. 19, no. 2, pp. 7–13, 2017.

[7] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, July 1948.

[8] W. Ryan and S. Lin, *Channel codes: classical and modern*. Cambridge university press, 2009.

[9] B. P. Smith, A. Farhood, A. Hunt, F. R. Kschischang, and J. Lodge, "Staircase codes: FEC for 100 Gb/s OTN," *IEEE J. Lightw. Technol.*, vol. 30, no. 1, pp. 110–117, Jan. 2012.

[10] L. M. Zhang and F. R. Kschischang, "Staircase codes with 6% to 33% overhead," *IEEE J. Lightw. Technol.*, vol. 32, no. 10, pp. 1999–2002, May 2014.

[11] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, and F. Forghieri, "The GN-model of fiber non-linear propagation and its applications," *IEEE J. Lightw. Technol.*, vol. 32, no. 4, pp. 694–721, Feb 2014.

[12] E. Torrengo, R. Cigliutti, G. Bosco, A. Carena, V. Curri, P. Poggiolini, A. Nespola, D. Zeolla, and F. Forghieri, "Experimental validation of an analytical model for nonlinear propagation in uncompensated optical links," *Opt. Express*, vol. 19, no. 26, pp. B790–B798, 2011.

[13] A. J. Stark, Y.-T. Hsueh, T. F. Detwiler, M. M. Filer, S. Tibuleac, and S. E. Ralph, "System performance prediction with the Gaussian noise model in 100G PDM-QPSK coherent optical networks," *IEEE J. Lightw. Technol.*, vol. 31, no. 21, pp. 3352–3360, 2013.

[14] M. Mushkin and I. Bar-David, "Capacity and coding for the Gilbert-Elliott channels," *IEEE Trans. Inform. Theory*, vol. 35, no. 6, pp. 1277–1290, 1989.

[15] T. Miya, Y. Terunuma, T. Hosaka, and T. Miyashita, "Ultimate low-loss single-mode fibre at 1.55 μm," *Electron. Lett.*, vol. 15, no. 4, pp. 106–108, 1979.

[16] G. Agrawal, *Nonlinear Fiber Optics, Fifth Edition.* Elsevier, 2013.

[17] D. Marcuse, C. Manyuk, and P. Wai, "Application of the Manakov-PMD equation to studies of signal propagation in optical fibers with randomly varying birefringence," *IEEE J. Lightw. Technol.*, vol. 15, no. 9, pp. 1735–1746, 1997.

[18] M. Karlsson, J. Brentel, and P. A. Andrekson, "Long-term measurement of PMD and polarization drift in installed fibers," *IEEE J. Lightw. Technol.*, vol. 18, no. 7, pp. 941–951, July 2000.

[19] G. Soliman, M. Reimer, and D. Yevick, "Measurement and simulation of polarization transients in dispersion compensation modules," *J. Opt. Soc. Am. A*, vol. 27, no. 12, pp. 2532–2541, 2010.

[20] P. M. Krummrich, D. Ronnenberg, W. Schairer, D. Wienold, F. Jenau, and M. Herrmann, "Demanding response time requirements on coherent receivers due to fast polarization rotations caused by lightning events," *Opt. Express*, vol. 24, no. 11, pp. 12 442–12 457, 2016.

[21] C. B. Czegledi, M. Karlsson, E. Agrell, and P. Johannisson, "Polarization drift channel model for coherent fibre-optic systems," *Scientific reports*, vol. 6, p. 21217, 2016.

[22] C. Poole and R. Wagner, "Phenomenological approach to polarisation dispersion in long single-mode fibres," *Electron. Lett.*, vol. 22, no. 19, pp. 1029–1030, 1986.

[23] G. Agrawal, *Fiber-Optic Communication Systems, Fouth Edition*, ser. Wiley Series in Microwave and Optical Engineering. Wiley, 2012.

[24] *EDFA100S*, https://www.thorlabs.com/newgrouppage9.cfm?objectgroup_ID=10680, Thorlabs, Inc., 2019, accessed: 2019-06-27.

[25] *EDFA-C-R*, https://oequest.com/getDatasheet/id/10817-10817.pdf, Optilab, LLC, 2019, accessed: 2019-06-27.

[26] A. Carena, G. Bosco, V. Curri, P. Poggiolini, M. T. Taiba, and F. Forghieri, "Statistical characterization of PM-QPSK signals after propagation in uncompensated fiber links," in *36th European Conference and Exhibition on Optical Communication*. IEEE, 2010, p. P.4.07.

[27] F. Vacondio, C. Simonneau, L. Lorcy, J. Antona, A. Bononi, and S. Bigo, "Experimental characterization of Gaussian-distributed nonlinear distortions," in *2011 37th European Conference and Exhibition on Optical Communication*. IEEE, 2011, p. We.7.B.1.

[28] D. E. Crivelli, M. R. Hueda, H. S. Carrer, M. del Barco, R. R. López, P. Gianni, J. Finochietto, N. Swenson, P. Voois, and O. E. Agazzi, "Architecture of a single-chip 50 Gb/s DP-QPSK/BPSK transceiver with electronic dispersion compensation for coherent optical channels," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 61, no. 4, pp. 1012–1025, April 2014.

[29] M. Tur, B. Moslehi, and J. Goodman, "Theory of laser phase noise in recirculating fiber-optic delay lines," *Journal of lightwave technology*, vol. 3, no. 1, pp. 20–31, 1985.

[30] L. Kull, T. Toifl, M. Schmatz, P. A. Francese, C. Menolfi, M. Braendli, M. Kossel, T. Morf, T. M. Andersen, and Y. Leblebici, "A 90GS/s 8b 667mW 64× interleaved SAR ADC in 32nm digital SOI CMOS," in *2014 IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC)*. IEEE, 2014, pp. 378–379.

[31] P. Serena, "Nonlinear signal-noise interaction in optical links with nonlinear equalization," *IEEE J. Lightw. Technol.*, vol. 34, no. 6, pp. 1476–1483, March 2016.

[32] H. J. Veendrick, *Nanometer CMOS ICs*. Springer, 2017.

[33] A. B. Kahng and S. Reda, "Intrinsic shortest path length: a new, accurate a priori wirelength estimator," in *Proceedings of the 2005 IEEE/ACM International conference on Computer-aided design*. IEEE Computer Society, 2005, pp. 173–180.

[34] D. Prasad, S. Sinha, B. Cline, S. Moore, and A. Naeemi, "Accurate processor-level wirelength distribution model for technology pathfinding using a modernized interpretation of Rent's rule," in *2018 55th*

ACM/ESDA/IEEE Design Automation Conference (DAC), June 2018, pp. 1–6.

[35] A. V. Oppenheim and C. J. Weinstein, "Effects of finite register length in digital filtering and the fast Fourier transform," *Proc. IEEE*, vol. 60, no. 8, pp. 957–976, Aug 1972.

[36] E. Ip and J. M. Kahn, "Compensation of dispersion and nonlinear impairments using digital backpropagation," *IEEE J. Lightw. Technol.*, vol. 26, no. 20, pp. 3416–3425, Oct 2008.

[37] Z. Tao, L. Dou, W. Yan, L. Li, T. Hoshida, and J. C. Rasmussen, "Multiplier-free intrachannel nonlinearity compensating algorithm operating at symbol rate," *IEEE J. Lightw. Technol.*, vol. 29, no. 17, pp. 2570–2576, 2011.

[38] L. Liu, L. Li, Y. Huang, K. Cui, Q. Xiong, F. N. Hauske, C. Xie, and Y. Cai, "Intrachannel nonlinearity compensation by inverse Volterra series transfer function," *IEEE J. Lightw. Technol.*, vol. 30, no. 3, pp. 310–316, 2011.

[39] A. Leven, N. Kaneda, U. Koc, and Y. Chen, "Frequency estimation in intradyne reception," *IEEE Photon. Technol. Lett.*, vol. 19, no. 6, pp. 366–368, March 2007.

[40] S. Zhang, L. Xu, J. Yu, M.-F. Huang, P. Y. Kam, C. Yu, and T. Wang, "Novel ultra wide-range frequency offset estimation for digital coherent optical receiver," in *Optical Fiber Communication Conference*. Optical Society of America, 2010, p. OWV3.

[41] M. Selmi, Y. Jaouen, and P. Ciblat, "Accurate digital frequency offset estimator for coherent polmux QAM transmission systems," in *2009 35th European Conference on Optical Communication*, Sep. 2009, p. P3.08.

[42] K. Mueller and M. Muller, "Timing recovery in digital synchronous data receivers," *IEEE transactions on communications*, vol. 24, no. 5, pp. 516–531, 1976.

[43] F. Gardner, "A BPSK/QPSK timing-error detector for sampled receivers," *IEEE Trans. Commun.*, vol. 34, no. 5, pp. 423–429, May 1986.

[44] B. Baeuerle, A. Josten, M. Eppenberger, D. Hillerkuss, and J. Leuthold, "Low-complexity real-time receiver for coherent Nyquist-FDM signals," *IEEE J. Lightw. Technol.*, vol. 36, no. 24, pp. 5728–5737, Dec 2018.

[45] M. Paskov, D. Lavery, and S. J. Savory, "Blind equalization of receiver in-phase/quadrature skew in the presence of Nyquist filtering," *IEEE Photon. Technol. Lett.*, vol. 25, no. 24, pp. 2446–2449, Dec 2013.

[46] D. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Transactions on Communications*, vol. 28, no. 11, pp. 1867–1875, Nov 1980.

[47] M. J. Ready and R. P. Gooch, "Blind equalization based on radius directed adaptation," in *International Conference on Acoustics, Speech, and Signal Processing*. IEEE, 1990, pp. 1699–1702.

[48] J. Mazo, "Analysis of decision-directed equalizer convergence," *Bell System Technical Journal*, vol. 59, no. 10, pp. 1857–1876, 1980.

[49] M. Mazur, J. Schröder, A. Lorences-Riesgo, M. Karlsson, and P. A. Andrekson, "Optimization of low-complexity pilot-based DSP for high spectral efficiency51 × 24Gbaud PM-64QAM transmission," in *2018 European Conference on Optical Communication (ECOC)*, Sep. 2018, pp. 1–3.

[50] A. Viterbi, "Nonlinear estimation of PSK-modulated carrier phase with application to burst digital transmission," *IEEE Transactions on Information theory*, vol. 29, no. 4, pp. 543–551, 1983.

[51] T. Pfau, S. Hoffmann, and R. Noe, "Hardware-efficient coherent digital receiver concept with feedforward carrier recovery for m-QAM constellations," *IEEE J. Lightw. Technol.*, vol. 27, no. 8, pp. 989–999, April 2009.

[52] E. Börjeson, C. Fougstedt, and P. Larsson-Edefors, "ASIC design exploration of phase recovery algorithms for M-QAM fiber-optic systems," in *Optical Fiber Communication Conference (OFC) 2019*. Optical Society of America, 2019, p. W3H.7.

[53] M. Renfors and Y. Neuvo, "The maximum sampling rate of digital filters under hardware speed constraints," *IEEE Trans. Circuits Syst.*, vol. 28, no. 3, pp. 196–202, Mar 1981.

[54] J. Chen and X. Liu, "A high-performance deeply pipelined architecture for elementary transcendental function evaluation," in *2017 IEEE International Conference on Computer Design (ICCD)*. IEEE, 2017, pp. 209–216.

[55] J. E. Volder, "The CORDIC trigonometric computing technique," *IRE Trans. Elec. Comp.*, no. 3, pp. 330–334, 1959.

[56] H. Kim, B.-G. Nam, J.-H. Sohn, J.-H. Woo, and H.-J. Yoo, "A 231-MHz, 2.18-mW 32-bit logarithmic arithmetic unit for fixed-point 3-D graphics system," *IEEE journal of solid-state circuits*, vol. 41, no. 11, pp. 2373–2381, 2006.

[57] R. W. Hamming, "Error detecting and error correcting codes," *The Bell system technical journal*, vol. 29, no. 2, pp. 147–160, 1950.

[58] R. Bose and D. Ray-Chaudhuri, "On a class of error correcting binary group codes," *Information and Control*, vol. 3, no. 1, pp. 68 – 79, 1960.

[59] A. Hocquenghem, "Codes correcteurs d'erreurs," *Chiffres*, vol. 2, no. 2, pp. 147–56, 1959.

[60] I. S. Reed and G. Solomon, "Polynomial codes over certain finite fields," *Journal of the society for industrial and applied mathematics*, vol. 8, no. 2, pp. 300–304, 1960.

[61] P. Elias, "Error-free coding," *IRE Trans. Inf. Theory*, vol. 4, no. 4, pp. 29–37, Sept. 1954.

[62] G. D. Forney, "Concatenated codes," *MIT Press*, 1965.

[63] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes. 1," in *Proceedings of ICC'93-IEEE International Conference on Communications*, vol. 2. IEEE, 1993, pp. 1064–1070.

[64] R. Gallager, "Low-density parity-check codes," *IRE Trans. Inf. Theory*, vol. 8, no. 1, pp. 21–28, 1962.

[65] R. J. Mears, L. Reekie, I. M. Jauncey, and D. N. Payne, "Low-noise erbium-doped fibre amplifier operating at 1.54 μm," *Electronics Letters*, vol. 23, no. 19, pp. 1026–1028, September 1987.

[66] H. Sun, K.-T. Wu, and K. Roberts, "Real-time measurements of a 40 Gb/s coherent system," *Opt. Express*, vol. 16, no. 2, pp. 873–879, Jan 2008.

[67] Y. Loussouarn, E. Pincemin, M. Pan, G. Miller, A. Gibbemeyer, and B. Mikkelsen, "Multi-rate multi-format CFP/CFP2 digital coherent interfaces for data center interconnects, metro, and long-haul optical communications," *IEEE J. Lightw. Technol.*, vol. 37, no. 2, pp. 538–547, Jan 2019.

[68] M. Nowell, A. Aranyosi, V. Le, J. J. Maki, S. Sommers, T. Palkert, and W. Chen, *QSFP-DD: Enabling 15 Watt Cooling Solutions*, http://www.qsfp-dd.com/wp-content/uploads/2018/03/QSFP-DD-Thermal-Whitepaper_Rev1_31018.pdf, Last accessed on April 14, 2019.

[69] B. Park, W. Meggitt, A. Bechtolsheim, C. Cole, B. Kirk, and C. Metivier, *OSFP octal small form factor pluggable module*, https://osfpmsa.org/assets/pdf/OSFP_Module_Specification_Rev2_0.pdf, Last accessed on June 24, 2019.

[70] J. Geyer, C. Rasmussen, B. Shah, T. Nielsen, and M. Givehchi, "Power efficient coherent transceivers," in *ECOC 2016; 42nd European Conference on Optical Communication*. VDE, 2016, pp. 1–3.

[71] C. Rasmussen, Y. Pan, M. Aydinlik, M. Crowley, J. Geyer, P. Humblet, F. Liu, B. Mikkelsen, P. Monsen, N. Nadarajah *et al.*, "Real-time DSP for 100+ Gb/s," in *Optical Fiber Communication Conference*. Optical Society of America, 2013, pp. OW1E–1.

[72] D. A. Morero, M. A. Castrillón, A. Aguirre, M. R. Hueda, and O. E. Agazzi, "Design tradeoffs and challenges in practical coherent optical transceiver implementations," *IEEE J. Lightw. Technol.*, vol. 34, no. 1, pp. 121–136, Jan 2016.

[73] M. Kuschnerov, T. Bex, and P. Kainzmaier, "Energy efficient digital signal processing," in *Optical Fiber Communication Conference*. Optical Society of America, 2014, pp. Th3E–7.

[74] L. Lundberg, P. A. Andrekson, and M. Karlsson, "Power consumption analysis of hybrid EDFA/Raman amplifiers in long-haul transmission systems," *IEEE J. Lightw. Technol.*, vol. 35, no. 11, pp. 2132–2142, 2017.

[75] J.-P. Elbers, N. Eiselt, A. Dochhan, D. Rafique, and H. Grießer, "PAM4 vs coherent for DCI applications," in *Signal Processing in Photonic Communications*. Optical Society of America, 2017, pp. SpTh2D–1.

[76] X. Zhou, R. Urata, and H. Liu, "Beyond 1Tb/s datacenter interconnect technology: Challenges and solutions," in *Optical Fiber Communication Conference*. Optical Society of America, 2019, pp. Tu2F–5.

[77] J. K. Perin, A. Shastri, and J. M. Kahn, "Design of low-power DSP-free coherent receivers for data center links," *IEEE J. Lightw. Technol.*, vol. 35, no. 21, pp. 4650–4662, Nov 2017.

[78] M. Morsy-Osman, M. Sowailem, E. El-Fiky, T. Goodwill, T. Hoang, S. Lessard, and D. V. Plant, "DSP-free "coherent-lite" transceiver for next generation single wavelength optical intra-datacenter interconnects," *Opt. Express*, vol. 26, no. 7, pp. 8890–8903, 2018.

[79] C. Bae, M. Gokhale, O. Gustafsson, and M. Garrido, "Improved implementation approaches for 512-tap 60 GSa/s chromatic dispersion FIR filters," in *Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, CA, USA, Oct. 2018, pp. 213–217.

[80] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Inter-channel nonlinear interference noise in WDM systems: modeling and mitigation," *IEEE J. Lightw. Technol.*, vol. 33, no. 5, pp. 1044–1053, 2014.

[81] O. Golani, M. Feder, and M. Shtaif, "NLIN mitigation using turbo equalization and an extended Kalman smoother," *IEEE J. Lightw. Technol.*, vol. 37, no. 9, pp. 1885–1892, May 2019.

[82] A. Sheikh, A. Graell i Amat, G. Liva, C. Hager, and H. D. Pfister, "On low-complexity decoding of product codes for high-throughput fiber-optic systems," in *2018 IEEE 10th International Symposium on Turbo Codes Iterative Information Processing (ISTC)*, Dec 2018, pp. 1–5.

[83] A. Sheikh, A. Graell i Amat, and G. Liva, "Binary message passing decoding of product codes based on generalized minimum distance decoding : (invited paper)," in *2019 53rd Annual Conference on Information Sciences and Systems (CISS)*, March 2019, pp. 1–5.

[84] Y. Lei, A. Alvarado, B. Chen, X. Deng, Z. Cao, J. Li, and K. Xu, "Decoding staircase codes with marked bits," in *2018 IEEE 10th International Symposium on Turbo Codes Iterative Information Processing (ISTC)*, Dec 2018, pp. 1–5.